VIROLOGY 233, 43–50 (1997)
ARTICLE NO. VY978615

Mutation Rate of GB Virus C/Hepatitis G Virus over the Entire Genome

View metadata, citation and similar papers at core.ac.uk
brought
provided by Elsevier

Haruhisa Nakao,* Hiroaki Okamoto,† Masako Fukuda,‡ Fumio Tsuda,§ Takehiro Mitsui,¶ Kazuo Masuko,¶ Hisao Iizuka,‖ Yuzo Miyakawa,** and Makoto Mayumi†,[1]

*First Department of Internal Medicine, Medical School, Nagoya City University, Aichi-Ken 467; †Immunology Division, Jichi Medical School, Tochigi-Ken 329-04; ‡Institute of Immunology, Tokyo 112; §Department of Medical Sciences, Toshiba General Hospital, Tokyo 140; ¶Masuko Hospital and Masuko Institute for Medical Research, Aichi-Ken 453; ‖Japanese Red Cross Saitama Blood Center, Saitama-Ken 338; and **Miyakawa Memorial Research Foundation, Tokyo 107, Japan

Received December 30, 1996; returned to author for revision February 13, 1997; accepted April 22, 1997

A patient on maintenance hemodialysis was infected with a recently discovered putative non-A to -E hepatitis virus designated GB virus C (GBV-C) or hepatitis G virus (HGV) by transfusion. The viral isolate was recovered from the patient soon after she turned positive for GBV-C/HGV RNA in serum (GSI85) and 8.4 years thereafter (GSI93), and the entire nucleotide sequences were determined. They both had a genomic length of 9391 nucleotides with a defective C gene made of only 42 nucleotides. Between GSI85 and GSI93, 31 (0.33%) nucleotides were different, which changed 5 (0.18%) of the encoded 2842 amino acids. Thus, GBV-C/HGV was estimated to have a mutation rate of 3.9×10^{-4} base substitutions per site per year. Nucleotide conversions were distributed over subgenomic regions, except in the 5′ untranslated region of 552 nucleotides and a defective short C gene, which were conserved in sequence. The change in the putative envelope genes (E1 and E2) was no different from that in the entire genome with only 6 (0.35%) nucleotide substitutions among the 1730, just 1 of which induced an amino acid conversion. Taken along with the comparison of the two isolates with the reported five GBV-C or HGV isolates, these results indicate that GBV-C/HGV would not have hypervariable regions and would use a strategy for viral persistence that is different from immune escape. © 1997 Academic Press

INTRODUCTION

The discovery of hepatitis C virus (HCV) has revealed it to be the major etiologic agent of blood-borne, acute as well as chronic, non-A, non-B hepatitis worldwide (Choo *et al.,* 1989; Kuo *et al.,* 1989). The exclusion of blood units with markers of HCV infection has decreased the risk of posttransfusion HCV infection virtually to zero (Holland, 1996; Schreiber *et al.,* 1996). The recognition of HCV infection, however, has unearthed a category of acute and chronic hepatitis without serological markers of known hepatitis viruses, and therefore is referred to as non-A to -E hepatitis (Alter and Bradley, 1995).

Recently, putative agents responsible for non-A to -E hepatitis have been cloned independently by two groups of investigators and their sequences determined (Leary *et al.,* 1996; Linnen *et al.,* 1996; Simons *et al.,* 1995). They are designated GB virus C (GBV-C) and hepatitis G virus (HGV). They both are a positive single-stranded RNA virus having sequence and genetic organization resembling *Flaviviridae* and are distantly related to HCV with a sequence divergence too wide to be classified as genotypes of HCV. GBV-C and HGV share 86% of nucleotide (nt) sequence and 96% of deduced amino acid (aa) sequence, and therefore are considered to be the same virus. Five genomes of GBV-C/HGV have been sequenced in full, including two from the United States, one from Africa, and two from Japan (Leary *et al.,* 1996; Linnen *et al.,* 1996; Okamoto *et al.,* 1997). Very recently, another GBV-C/HGV genome has been sequenced in its entirety (Erker *et al.,* 1996). They may be classified into three distinct genotypes provisionally designated G1, G2, and G3 (Okamoto *et al.,* 1997). Muerhoff *et al.* (1996) also have found evidence for three major groups of GBV-C/HGV based on the sequence heterogeneity within a 5′-terminal sequence of 600 bp.

The five GBV-C/HGV genomes have a similarity in the nucleotide sequence of >86% which far exceeds that of reported HCV genomes at >65% (Bukh *et al.,* 1995; Miyakawa *et al.,* 1995; Simmonds, 1995). Therefore, GBV-C/HGV may have a mutation rate lower than the 1.4–1.9 $\times 10^{-3}$ base substitutions per site per year reported for HCV (Ogata *et al.,* 1991; Okamoto *et al.,* 1992a). We have cloned and sequenced GBV-C/HGV isolates recovered from a patient on hemodialysis at an interval of 8.4 years and compared them within the entire genome and subgenomic regions.

The nucleotide sequence data reported in this article have been deposited with the DDBJ, EMBL, and GenBank Databases under Accession No. D87262 for the GSI85 isolate and D87263 for the GSI93 isolate.

[1] To whom correspondence and reprint requests should be addressed at Immunology Division, Jichi Medical School, Minamikawachi-Machi, Tochigi-Ken 329-04, Japan. Fax: (81) 285-44-1557.

0042-6822/97 $25.00
Copyright © 1997 by Academic Press
All rights of reproduction in any form reserved.

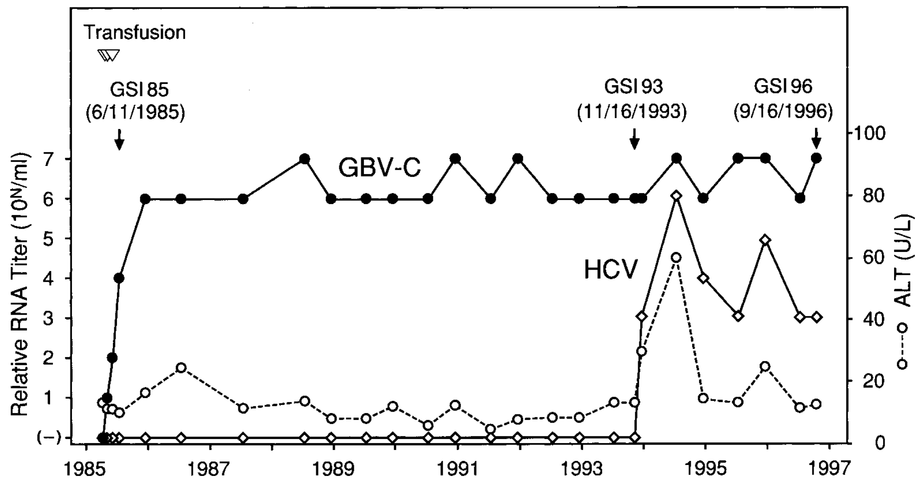

FIG. 1. Clinical course of a patient on maintenance hemodialysis who contracted infection with GB virus C/hepatitis G virus by transfusion. Relative titers of GB virus C/hepatitis G virus and fluctuating levels of alanine aminotransferase (ALT) are shown. Triangles represent transfusions, and arrows indicate blood samplings from which GSI85, GSI93, and GSI96 isolates were recovered.

MATERIALS AND METHODS

Serum samples

A patient (female, 45 years) with chronic renal failure received transfusion with one unit of concentrated erythrocytes (recovered from 200 ml of blood) on March 28 and April 8 and then two units on May 16, 1985 [patient 14 reported by Masuko *et al.* (1996)]. She was negative for GBV-C/HGV RNA in serum taken on March 26, and found to have turned positive on April 30 and remained so thereafter (Fig. 1). She was found to be infected with HCV on December 28, 1993, and seroconverted to antibody to HCV in February 1994. The antibody gradually increased to a high titer (absorbance in enzyme immunoassay >2.00) in May 1994 and thereafter. GBV-C/HGV isolate in her serum taken on June 11, 1985, when the viral RNA was increasing in titer, and that in the last serum harvested on November 16, 1993, before she turned positive for HCV RNA, were amplified by polymerase chain reaction (PCR) to determine the entire nucleotide sequences. They were designated GSI85 and GSI93, respectively, representing two GBV-C/HGV isolates recovered from the patient with an interval of 8.4 years. Her alanine aminotransferase levels stayed within normal limits (<45 U/L) through the 8.4 years, but were elevated after she contracted HCV infection in parallel with HCV RNA titers. The envelope genes of GBV-C/HGV (E1 and E2), as well as the hypervariable region (HVR) of HCV spanning 25 aa in the E2 gene (Hijikata *et al.,* 1991; Ogata *et al.,* 1991; Okamoto *et al.,* 1992a; Weiner *et al.,* 1991), were sequenced and compared at an interval of 2.8 years.

Semiquantification of GBV-C/HGV RNA

Total RNAs were extracted from serum (100 μl) using the extraction reagent containing guanidinium isothiocyanate and phenol (ISOGEN-LS from Nippon Gene Co, Ltd., Tokyo, Japan) and dissolved in 5.3 μl of distilled water pretreated with diethylpyrocarbonate. They were heated at 70° for 1 min, chilled quickly on ice, converted to cDNA, and tested for GBV-C/HGV RNA by PCR with primers deduced from the 5′ untranslated region (UTR) by the method described previously (Shimizu *et al.,* 1996). Serial 10-fold dilutions of extracted RNA were tested, and the relative concentration of GBV-C/HGV RNA was expressed by the reciprocal of the highest dilution in which the viral RNA was detectable; it was converted to represent the titer (10^N) per milliliter.

Determination of the genomic sequence of GBV-C/HGV

Nucleic acids extracted from 100 μl of serum were dissolved in distilled RNase-free water and denatured at 70° for 1–3 min. They were transcribed into cDNA with antisense primers and cloned reverse transcriptase of Molony murine leukemia virus (Superscript II; GIBCO BRL, Gaithersburg, MD). cDNA was heated at 95° for 15 min and subjected to PCR with TaKaRa Ex *Taq* polymerase (TaKaRa Biochemicals, Kyoto, Japan) for 35 cycles (94°, 30 sec; 55°, 30 sec; 72°, 45–150 sec [8 min in the last cycle]). One-tenth of the product was amplified by a second PCR for 25–35 cycles under the same conditions.

Strategy of sequencing GBV-C/HGV is shown in Fig. 2. GSI85 was amplified in six domains (regions *a* to *f*) and GSI93 in seven (*c, d,* and *g* to *k*). Either conserved or isolate-dependent sense and antisense primers were used to amplify regions *b* to *e* and *h* to *j* (sequences shown in the legend to Fig. 2), while a single-sided PCR method was used to amplify the other regions. cDNAs corresponding to the 5′-terminal sequences (regions *a* and *g*) were tailed with dATP or dTTP homopolymer at the 3′-end by terminal deoxynucleotidyl transferase (Boehringer Mannheim, Mann-

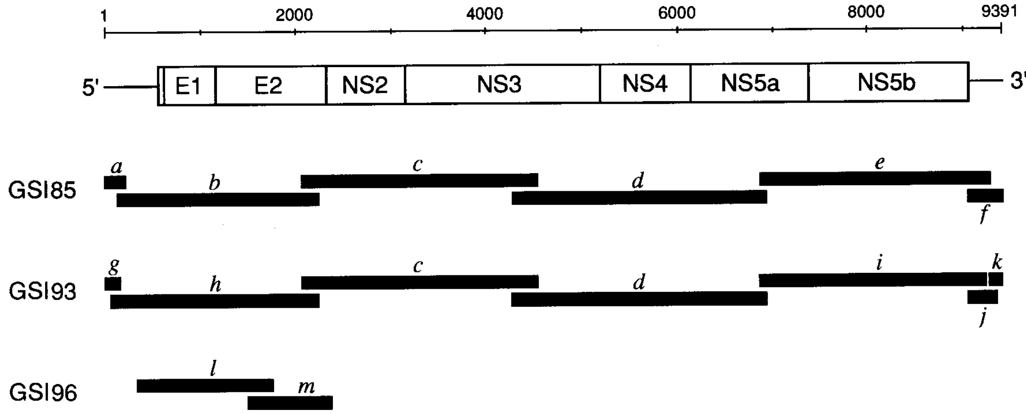

FIG. 2. The genomic organization and strategy for sequencing GSI85, GSI93, and GSI96 isolates. The nucleotide sequence was numbered from the putative 5′-end, and the genomic organization was estimated from open reading frames recognized in the genomic sequence. Closed bars represent cDNAs amplified by PCR with appropriate primers. They spanned *a,* nt 1–190; *b,* nt 130–2234; *c,* nt 2175–4436; *d,* nt 4279–6933; *e,* nt 6883–9296; *f,* nt 9044–9391; *g,* nt 1–151; *h,* nt 43–2234; *i,* nt 6883–9251; *j,* nt 9044–9351; *k,* nt 9281–9391; *l,* nt 338–1760; and *m,* nt 1493–2371. Primers and their sequences for amplification of each genomic region were *a,* 166 (5′-AGG GAT CCG TCG ACA TCG AT-3′) and G111 (5′-CCY TTW AGA CCC ACC TAT AGT-3′); *b,* G58 (5′-CAG GGT TGG TAG GTC GTA AAT CC-3′) and G82 (5′-GCC TCA GCC AGC TTC ATC AG-3′); *c,* G25 (5′-TCG CTG GTT GCT GCT GGA CTT TGT-3′) and G9 (5′-TCY TTG ATG ATD GAA CTG TC-3′); *d,* G8 (5′-TAT GGG CAT GGH ATH CCY CT-3′) and G47 (5′-AGG CAT CTC TCG GCT ACT TCC ACA-3′); *e,* G137 (5′-ATG GAG GAC TGC AGT ACA CC-3′) and G129 (5′-CCG TAG TCA CGG ATT ACG CT-3′); *f,* G126 (5′-GGT TCT TAG CCC TGC TCA TC-3′) and 167 (5′-CCG TCG ACA TCG ATA ATA CG-3′); *g,* 166 (5′-AAG GAT CCG TCG ACA TCG AT-3′) and G123 (5′-GAT TTA AGA TCT ACC AAC CCT G-3′); *h,* G71 (5′-AGC CCC AGA AAC CGA CGC CTA T-3′) and G82 (5′-GCC TCA GCC AGC TTC ATC AG-3′); *i,* G137 (5′-ATG GAG GAC TGC AGT ACA CC-3′) and G128 (5′-GCA GTG CAT TAG CAC CAT CA-3′); *j,* G126 (5′-GGT TCT TAG CCC TGC TCA TC-3′) and G130 (5′-ACC CCT TCA GAT CAC AGT GC-3′); *k,* G84 (5′-TAA TCC GTG ACT ACG GGC TG-3′) and 167 (5′-CCG TCG ACA TCG ATA ATA CG-3′); *l,* G36 (5′-CGT CGC CCT TCA ATG TCT CTC T-3′) and G172 (5′-CTR TTG RCG AAG GGC ACA GC-3′); and *m,* G171 (5′-CGA AGA TCG ATG TGT GGA GT-3′) and G33 (5′-GGA GCC CAA GAC ACC AGG ACA AGG-3′). Ambiguity codes are Y = T or C; W = A or T; R = A or G; D = G, A, or T; and H = A, T, or C.

heim, Germany). They were amplified by PCR with 5′UTR primers (antisense) and 43-mer oligonucleotides containing $(T)_{17}$ or $(A)_{17}$; these have been described as primer 171 or 165 (Okamoto *et al.,* 1992b). In order to determine the extreme 3′-end sequences (regions *f* and *k*), extracted RNAs were tailed with poly(A) by poly(A) polymerase (TaKaRa Biochemicals), converted to cDNA with the 43-mer oligonucleotide containing $(T)_{17}$, and subjected to amplification by a single-sided PCR.

Amplification products were separated on gel electrophoresis and ligated into M13 phage vector, and cDNA clones representing each region were obtained by the method described previously (Okamoto *et al.,* 1992b). Both plus and minus strands were sequenced by ALF AutoRead DNA sequencing kit (Pharmacia LKB Biotechnology, Uppsala, Sweden), and the consensus sequence of three clones was adopted for each region.

The envelope genes of a GBV-C/HGV isolate (GSI96), recovered 2.8 years after the patient became infected with HCV, were sequenced over two regions (E1 and E2) after amplification by PCR with primers deduced from conserved areas.

Sequencing the HVR of HCV

The patient contracted infection with HCV of genotype II/1b. A sequence of 146 bp including HVR (nt 1467–1612) was amplified by PCR with primers matching the sequence of genotype II/1b, 412 (sense, 5′-CTA CTC TTC GCC GGC GTT GA-3′ [nt 1467–1486]) and 413 (antisense, 5′-CTG TTG ATG TGC CAG CTG CC-3′ [nt 1593–1612]). Ten clones each were propagated from sera obtained at two time points and sequenced.

Computer analysis

Sequence analysis was performed using ODEN Version 1.1.1 (National Institute of Genetics, Mishima, Japan) and GENETYX Version 8.0 (Software Development Co., Ltd., Tokyo, Japan) programs. A phylogenetic tree of GBV-C/HGV was constructed by the unweighted pair-group method with arithmetic mean (Nei, 1987) and the neighbor-joining method (Saitou and Nei, 1987).

RESULTS

GSI85 and GSI93 isolates of GBV-C/HGV

GSI85 isolate recovered from a hemodialysis patient soon after she contracted GBV-C/HGV infection and GSI93 isolate obtained from her 8.4 years later (Fig. 1) were sequenced in full. They both consisted of 9391 bases and contained a long open reading frame (ORF) spanning nt 553 to 9078 and coding for 2842 aa flanked by putative 5′UTR (nt 1 to 552) and 3′UTR (nt 9079 to 9391) that did not have a poly(U) stretch or poly(A) tail. Their ORF was divided into the C gene which was defective and encoded only 14 aa, putative E1 gene (190 aa) and E2 gene (387 aa), as well as NS2 (281 aa), NS3 (677

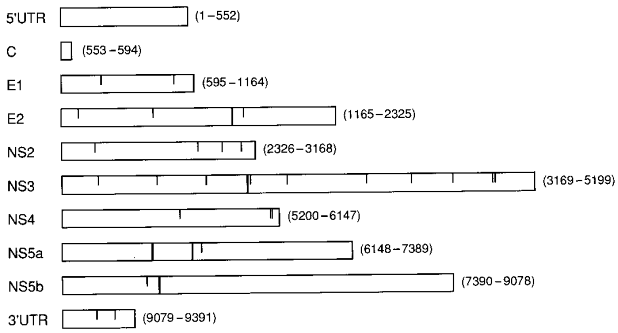

FIG. 3. Nucleotide differences between GSI85 and GSI93 isolates. Each genomic region is represented by a box followed by the sequence position in parentheses. Silent mutations without amino acid changes are indicated by short thin bars and missense mutations by long thick bars.

aa), NS4 (316 aa), NS5a (414 aa), and NS5b (563 aa) regions. These genes and regions were deduced by comparison with those of GB viruses A to C, HGV, and HCV (Grakoui *et al.,* 1993; Leary *et al.,* 1996; Linnen *et al.,* 1996; Muerhoff *et al.,* 1995; Okamoto *et al.,* 1997).

GSI85 and GSI93 were different in 31 (0.3%) of the 9391 bases. Based on this difference, the mutation rate of GBV-C/HGV was estimated to be 3.9×10^{-4} base substitutions per site per year. Of 29 nucleotide conversions in the ORF, 5 (17%) were in the first letter, 3 (11%) in the second letter, and the remaining 21 in the third letter; they induced only five amino acid changes. The distribution of missense and silent mutations is shown in Fig. 3. They are scattered more or less evenly over various subgenomic areas. Remarkably, mutations were not detected in 5′UTR and the 3′-terminal two thirds of NS5b. They were not seen in 42 bases for the defective C gene of GSI85 and GSI93, either. Only one of the five missense mutations was in the envelope genes (E2). In addition, there was one missense mutation in NS3, two in NS5a, and one in NS5b.

Of the 31 mutations observed in GSI93, 1 was detected already in one of the three clones from GSI85. Since the mutation occurred in 3′UTR, it did not accompany an amino acid change. Another 2 nonsense mutations were found in two of the three clones in GSI93. The remaining clone had the same nucleotide as the one unanimously found in all three clones from GSI85.

The sequence and length of the extreme 3′-end was the same for all three clones from GSI93 and GSI85. The extreme 5′-end in some clones was shorter than in others. The longest ones with nt 1 being T as in GT230 and GT110 (Okamoto *et al.,* 1997) were considered to cover the extreme 5′-end, and their sequences were adopted to represent GSI93 and GSI85.

Comparison of GSI85 and GSI93 isolates with five GBV-C/HGV genomes of which the full sequence is known

Table 1 compares GSI93 with the five reported GBV-C/HGV genomes, as well as with GSI85, within the total genome and subgenomic regions. GSI93 was closest to GT230 reported from Japan with an overall similarity of 91.6% in nucleotide sequence and 98.6% in deduced amino acid sequence. It was similar to the other four genomes by 87.1–87.2% in nucleotide sequence and 96.6–97.6% in amino acid sequence. There were no subgenomic regions that diverged much more than the others. Although GSI93 was similar to HGV-R10291 only in 64.3% of amino acid sequence of the defective C protein, only 14 amino acids for a C-gene product of GSI93 were compared. A low similarity of 53.1% was observed in 3′UTR between GSI93 and prototype GBV-C; the comparison was made on merely 64 overlapping nucleotides.

These results indicated that GSI93 and GSI85 would belong to the same genotype of GBV-C/HGV tentatively designated G3. In Fig. 4, a phylogenetic tree was constructed on seven GBV-C/HGV isolates, including GSI93 as well as GSI85 and the five reported, within an overlapping sequence of 8526 bases in the ORF; it corresponded to 91% of the longest genomic sequence displayed by GT110. The tree was built by the unweighted pair-group method with arithmetic mean (Nei, 1987), and revealed three genotypes of GBV-C/HGV; it indicated GSI85 and GSI93 as members of genotype G3. The three genotypes were confirmed by another tree constructed by the neighbor-joining method (Saitou and Nei, 1987).

Changes in the amino acid sequence of the putative envelope proteins of GBV-C/HGV and those in HVR of co-infecting HCV

Table 2 lists nucleotide mutations and induced amino acid conversions within the E1 and E2 genes of GBV-C/HGV isolates during 8.4 years and an additional 2.8 years after the patient became infected with HCV. Nine point mutations occurred during 11.2 years with only two amino acid conversions in GBV-C/HGV, with an estimated mutation rate of 4.6×10^{-4} nucleotide substitutions per site per year.

The sequence of 25 aa representing HVR is shown in Fig. 5 for HCV recovered at the first detection of HCV RNA and 2.8 years thereafter. Of 10 clones at the first detection, 5 (clones 1–5) had the identical sequence and another (clone 6) differed from them in only 1 aa. The remaining 4 (clones 7–10) had the same sequence which possessed 11–12 aa conversions from the others.

None of the clones 11–20, recovered 2.8 years later, shared identical sequence of HVR with any clones from the first detection of HCV. One of them (clone 11) was deduced to have evolved from clones 1–6, differing in 3–4 aa, while all the others (clones 12–20) were from clones 7–10, differing in 7–9 aa. These amino acid conversions were induced by mutations in 3–4 nt and 8–11 nt, respectively. Thus a nucleotide substitution rate per

TABLE 1

Percentage of Similarity of Nucleotide and Deduced Amino Acid Sequences within the Entire Genome and Subgenomic Regions between GSI93 and Five Reported GB Virus C/Hepatitis G Virus Isolates as Well as GSI85[a]

Genomic region	Nucleotides (amino acids)	GBV-C/HGV genotypes					
		G1	G2			G3	
		GBV-C	HGV-PNF2161	HGV-R10291	GT110	GT230	GSI85
Entire	9103–9395	87.2	87.2	87.1	87.2	91.6	99.7
	(2842–2933)	(97.6)	(96.6)	(96.6)	(97.6)	(98.6)	(99.8)
5′UTR	275–552	89.5	90.6	91.6	91.4	96.0	100
C	42–315	88.1	88.1	81.0	88.1	88.1	100
	(14–105)	(92.9)	(92.9)	(64.3)	(92.9)	(92.9)	(100)
E1	570	89.1	85.4	85.1	84.7	90.5	99.6
	(190)	(97.9)	(96.8)	(95.8)	(96.8)	(98.4)	(100)
E2	1161	86.5	85.4	85.1	85.0	92.3	99.7
	(387)	(96.4)	(91.5)	(94.3)	(95.6)	(97.7)	(99.7)
NS2	843	84.7	83.9	85.8	83.4	89.2	99.5
	(281)	(97.5)	(95.7)	(96.8)	(97.2)	(98.2)	(100)
NS3	2031	85.9	86.9	87.5	86.1	88.5	99.5
	(677)	(98.1)	(98.5)	(98.7)	(99.0)	(99.1)	(99.9)
NS4	948	86.5	85.9	86.6	86.1	91.6	99.7
	(316)	(97.2)	(95.9)	(96.8)	(96.8)	(98.1)	(100)
NS5a	1242	89.9	88.0	87.8	88.1	93.3	99.8
	(414)	(98.8)	(96.9)	(96.4)	(97.3)	(99.0)	(99.5)
NS5b	1689–1692	88.8	88.3	88.2	89.6	92.2	99.9
	(563–564)	(97.3)	(98.2)	(98.0)	(98.6)	(99.1)	(99.8)
3′UTR	64–315	53.1	96.2	91.8	96.2	98.4	99.4

[a] Genomic length and accession numbers of five reported isolates are prototype GBV-C, 9125 nt/U36380; HGV-PNF2161, 9392nt/U44402; HGV-R10291, 9103 nt/U45966; GT110, 9395 nt/D90600; and GT230, 9390 nt/D90601.

site per year of 1.4×10^{-2} to 5.2×10^{-1} was estimated for HVR of HCV during 2.8 years.

DISCUSSION

GBV-C/HGV prevails across the world, infecting approximately 1–2% of apparently healthy blood donors in the United States (Linnen *et al.,* 1996), Japan (Masuko *et al.,* 1996), Italy (Fiordalisi *et al.,* 1996), and China (Wang *et al.,* 1997). Hence, GBV-C/HGV would be a common virus persistently infecting human beings. It transmits parenterally, typically by transfusion and intravenous drug abuse (Aikawa *et al.,* 1996; Linnen *et al.,* 1996; Masuko *et al.,* 1996; Schmidt *et al.,* 1996). GBV-C/HGV RNA is detected frequently in patients with chronic hepatitis

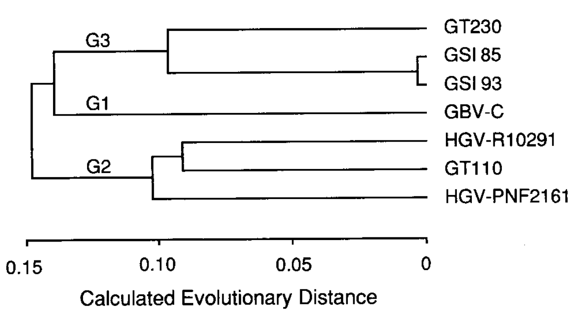

FIG. 4. A phylogenetic tree of GB virus C/hepatitis G virus and three proposed genotypes. The extreme 5′- and 3′-terminal sequences are not determined for two isolates (prototype GBV-C and HGV-R10291). Sequence was compared within 8526 bases spanning nt 553–9078 in the open reading frame, which have been determined for all of the seven GBV-C/HGV isolates.

TABLE 2

Nucleotide and Amino Acid Changes in the E1 Gene (nt 595–1164 [570 bp]) and E2 Gene (nt 1165–2325 [1161 bp]) of GB Virus C/Hepatitis G Virus during 8.4 Years and an Additional 2.8 Years

Nucleotide position	Date tested					
	6/11/1985		11/16/1993		9/16/1996	
E1 gene						
771 (aa 73)	C	(Pro)	*T*	(Pro)	T	(Pro)
864 (aa 104)	G	(Gly)	G	(Gly)	*A*	(Gly)
1077 (aa 175)	G	(Leu)	*A*	(Leu)	A	(Leu)
E2 gene						
1239 (aa 229)	T	(Asn)	*C*	(Asn)	C	(Asn)
1510 (aa 320)	A	(Ser)	A	(Ser)	*G*	(*Gly*)
1557 (aa 335)	C	(Gly)	*T*	(Gly)	T	(Gly)
1903 (aa 451)	A	(Lys)	*G*	(*Glu*)	G	(Glu)
1953 (aa 467)	T	(Asn)	*C*	(Asn)	C	(Asn)
2304 (aa 584)	G	(Pro)	G	(Pro)	*A*	(Pro)

[a] Mutated nucleotides and induced amino acid conversions are in italic.

At the First Detection (12/28/1993)

```
 1: STTTVGGSQASTTYRLTSLFTSGPS
 2: -------------------------
 3: -------------------------
 4: -------------------------
 5: -------------------------
 6: -------------------L-----
 7: ----I--TN-R--SGFVRF--L---
 8: ----I--TN-R--SGFVRF--L---
 9: ----I--TN-R--SGFVRF--L---
10: ----I--TN-R--SGFVRF--L---
```

After 2.8 Years (9/16/1996)

```
11: F-----------H--A--------
12: D-R-I--TS-R--SG-V----P---
13: D-R-I--TS-R--SG-V----P---
14: D-R-I--TS-R--SG-V----P---
15: D-R-I--TS-R--SG-V----P---
16: D-R-I--TS-R--SG-V----P---
17: D-R-I--TS-R--SG-V----P---
18: D-R-I--TS-R--SG-VG---P---
19: D-R-I--TS-R--SG-V----P--P
20: G-R----AS-R--SG-VG---P---
```

Amino Acids 384-408
of Hepatitis C Virus

FIG. 5. Amino acid sequences of the HVR of co-infecting HCV determined at the first recognition of infection and 2.8 years thereafter.

C, and is particularly common in patients on maintenance hemodialysis, with prevalence rates ranging from 3.1 to 58% (Alter, 1996; de Lamballerie *et al.,* 1996; Masuko *et al.,* 1996; Tsuda *et al.,* 1996; Wang *et al.,* 1997). Although most individuals infected with GBV-C/HGV have normal levels of serum transaminases, there are some patients with fulminant or chronic hepatitis of a non-A to -E etiology who are infected with it (Fiordalisi *et al.,* 1996; Heringlake *et al.,* 1996; Linnen *et al.,* 1996; Yoshiba *et al.,* 1995). However, the site of replication or hepatitis-inducing activity is, as yet, not clear for GBV-C/HGV.

Since GBV-C/HGV has not been propagated in culture, the nucleotide sequences of viral isolates recovered from infected individuals remain the only means of knowing its virological characteristics. Being an RNA virus that evolves rapidly (Holland *et al.,* 1982), GBV-C/HGV is expected to have a high mutation rate. We determined the entire sequence of two GBV-C/HGV isolates recovered from a patient on maintenance hemodialysis at an interval of 8.4 years, which were named GSI85 and GSI93 isolates, respectively. Both GSI85 and GSI93 had a genomic length of 9391 bases, and they differed in 31 bases (0.3%) of which only 5 induced amino acid changes among the 2842 amino acids, corresponding to only 0.2%. They were classified into genotype G3 along with GT230 isolate reported from Japan (Okamoto *et al.,* 1997).

The mutation rate estimated by comparison of GSI85 and GSI93 isolates was 3.9×10^{-4} base substitutions per site per year. Since they were both defective, lacking the C gene as in all reported isolates (Erker *et al.,* 1996; Leary *et al.,* 1996; Linnen *et al.,* 1996; Okamoto *et al.,* 1997), it is not certain if the estimate could be taken as that of the complete GBV-C/HGV genome. The obtained rate was less than that of hepatitis C virus ($1.4-1.9 \times 10^{-3}$) (Ogata *et al.,* 1991; Okamoto *et al.,* 1992a). Mutations for amino acid changes were induced by 38% of nucleotide substitutions in HCV, at a rate much higher than 16% found for GBV-C/HGV in the present study.

Like HCV, 5′UTR was most conserved during 8.4 years without any mutations in the 552 nt constituting it. Although 3′UTR was most variable, only 2 (0.6%) of 313 nt changed. Within the coding region, conversions of 29 nt, of which 5 induced amino acid changes, were scattered over the entire genome. No subgenomic regions were identified where mutations clustered. Remarkably, few mutations occurred in the putative envelope genes (E1/E2) with only one amino acid change.

In the HCV genome, there is an HVR in the N-terminus of the E2 protein spanning 25–30 amino acids (Hijikata *et al.,* 1991; Ogata *et al.,* 1991; Okamoto *et al.,* 1992a; Weiner *et al.,* 1991). Since GBV-C/HGV is devoid of HVR, it may use different means of escaping immune clearance by the host. It might be coated with lipoproteins, as HCV is (Miyamoto *et al.,* 1992; Prince *et al.,* 1996; Sato *et al.,* 1993; Thomssen *et al.,* 1992), or the other host moieties to prevent surface proteins from binding with antibodies. In support of this view, GBV-C/HGV is precipitated by antibodies to human apolipoprotein A-I or apolipoprotein B, but hardly by antibodies to human IgG (Sato *et al.,* 1996).

It would have to be pointed out that the mutation rate of GBV-C/HGV was estimated in a patient on maintenance hemodialysis who is known to have compromised immune responses (Descamps-Latscha and Herbelin, 1993; Goldblum and Reed, 1980). However, HVR of co-infecting HCV mutated with a nucleotide substitution per site per year of 1.4×10^{-2} to 5.2×10^{-1}, which is comparable to that in immune-competent individuals (Higashi *et al.,* 1993; Taniguchi *et al.,* 1993). Since the hemodialysis patient we studied could elicit antibody to HCV in due course, she would not have been totally immunocompromised. The low mutation rate observed in her, therefore, would not be ascribable to compromised immune responses inherent to hemodialysis patients.

REFERENCES

Aikawa, T., Sugai, Y., and Okamoto, H. (1996). Hepatitis G infection in drug abusers with chronic hepatitis C. *N. Engl. J. Med.* 334, 195–196. [Letter]

Alter, H. J. (1996). The cloning and clinical implications of HGV and HGBV-C. *N. Engl. J. Med.* 334, 1536–1537.

Alter, H. J., and Bradley, D. W. (1995). Non-A, non-B hepatitis unrelated to the hepatitis C virus (non-ABC). *Semin. Liver Dis.* 15, 110–120.

Bukh, J., Miller, R. H., and Purcell, R. H. (1995). Genetic heterogeneity of hepatitis C virus: Quasispecies and genotypes. *Semin. Liver Dis.* 15, 41–63.

Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., and Houghton, M. (1989). Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 244, 359–362.

Descamps-Latscha, B., and Herbelin, A. (1993). Long-term dialysis and cellular immunity: A critical survey. *Kidney Int.* 41, S135–142.

de Lamballerie, X., Charrel, R. N., and Dussol, B. (1996). Hepatitis GB virus C in patients on hemodialysis. *N. Engl. J. Med.* 334, 1549. [Letter]

Erker, J. C., Simons, J. N., Muerhoff, A. S., Leary, T. P., Chalmers, M. L., Desai, S. M., and Mushahwar, I. K. (1996). Molecular cloning and characterization of a GB virus C isolate from a patient with non-A–E hepatitis. *J. Gen. Virol.* 77, 2713–2720.

Fiordalisi, G., Zanella, I., Mantero, G., Bettinardi, A., Stellini, R., Paraninfo, G., Cadeo, G., and Primi, D. (1996). High prevalence of GB virus C infection in a group of Italian patients with hepatitis of unknown etiology. *J. Infect. Dis.* 174, 181–183.

Goldblum, S. E., and Reed, W. P. (1980). Host defenses and immunologic alterations associated with chronic hemodialysis. *Ann. Intern. Med.* 93, 597–613.

Grakoui, A., Wychowski, C., Lin, C., Feinstone, S. M., and Rice, C. M. (1993). Expression and identification of hepatitis C virus polyprotein cleavage products. *J. Virol.* 67, 1385–1395.

Heringlake, S., Osterkamp, S., Trautwein, C., Tillmann, H. L., Böker, K., Muerhoff, S., Mushawar, I. K., Hunsmann, G., and Manns, M. P. (1996). Association between fulminant hepatic failure and a strain of GBV virus C. *Lancet* 348, 1626–1629.

Higashi, Y., Kakumu, S., Yoshioka, K., Wakita, T., Mizokami, M., Ohba, K., Ito, Y., Ishikawa, T., Takayanagi, M., and Nagai, Y. (1993). Dynamics of genome change in the E2/NS1 region of hepatitis C virus *in vivo. Virology* 197, 659–668.

Hijikata, M., Kato, N., Ootsuyama, Y., Nakagawa, M., Ohkoshi, S., and Shimotohno, K. (1991). Hypervariable regions in the putative glycoprotein of hepatitis C virus. *Biochem. Biophys. Res. Commun.* 175, 220–228.

Holland, J., Spindler, K., Horodyski, F., Grabau, E., Nichol, S., and VandePol, S. (1982). Rapid evolution of RNA genomes. *Science* 215, 1577–1585.

Holland, P. V. (1996). Viral infections and the blood supply. *N. Engl. J. Med.* 334, 1734–1735.

Kuo, G., Choo, Q. L., Alter, H. J., Gitnick, G. L., Redeker, A. G., Purcell, R. H., Miyamura, T., Dienstag, J. L., Alter, M. J., Stevens, C. E., Tegtmeier, G. E., Bonino, F., Colombo, M., Lee, W. S., Kuo, C., Berger, K., Shuster, J. R., Overby, L. R., Bradley, D. W., and Houghton, M. (1989). An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* 244, 362–364.

Leary, T. P., Muerhoff, A. S., Simons, J. N., Pilot-Matias, T. J., Erker, J. C., Chalmers, M. L., Schlauder, G. G., Dawson, G. J., Desai, S. M., and Mushahwar, I. K. (1996). Sequence and genomic organization of GBV-C: A novel member of the Flaviviridae associated with human non-A–E hepatitis. *J. Med. Virol.* 48, 60–67.

Linnen, J., Wages, J., Zhang-Keck, Z. Y., Fry, K. E., Krawczynski, K. Z., Alter, H., Koonin, E., Gallagher, M., Alter, M., Hadziyannis, S., Karayiannis, P., Fung, K., Nakatsuji, Y., Shih, J. W. K., Young, L., Piatak, M., Hoover, C., Fernandez, J., Chen, S., Zou, J. C., Morris, T., Hyams, K. C., Ismay, S., Lifson, J. D., Hess, G., Foung, S. K. H., Thomas, H., Bradley, D., Margolis, H., and Kim, J. P. (1996). Molecular cloning and disease association of hepatitis G virus: A transfusion-transmissible agent. *Science* 271, 505–508.

Masuko, K., Mitsui, T., Iwano, K., Yamazaki, C., Okuda, K., Meguro, T., Murayama, N., Inoue, T., Tsuda, F., Okamoto, H., Miyakawa, Y., and Mayumi, M. (1996). Infection with hepatitis GB virus C in patients on maintenance hemodialysis. *N. Engl. J. Med.* 334, 1485–1490.

Miyakawa, Y., Okamoto, H., and Mayumi, M. (1995). Classifying hepatitis C virus genotypes. *Mol. Med. Today* 1, 20–25.

Miyamoto, H., Okamoto, H., Sato, K., Tanaka, T., and Mishiro, S. (1992). Extraordinarily low density of hepatitis C virus estimated by sucrose density gradient centrifugation and the polymerase chain reaction. *J. Gen. Virol.* 73, 715–718.

Muerhoff, A. S., Leary, T. P., Simons, J. N., Pilot-Matias, T. J., Dawson, G. J., Erker, J. C., Chalmers, M. L., Schlauder, G. G., Desai, S. M., and Mushahwar, I. K. (1995). Genomic organization of GB viruses A and B: Two new members of the Flaviviridae associated with GB agent hepatitis. *J. Virol.* 69, 5621–5630.

Muerhoff, A. S., Simons, J. N., Leary, T. P., Erker, J. C., Chalmers, M. L., Pilot-Matias, T. J., Dawson, G. J., Desai, S. M., and Mushahwar, I. K. (1996). Sequence heterogeneity within the 5′-terminal region of the hepatitis GB virus C genome and evidence for genotypes. *J. Hepatol.* 25, 379–384.

Nei, M. (1987). Phylogenetic trees. *In* “Molecular Evolutionary Genetics” (M. Nei, Ed.), pp. 287–326. Columbia Univ. Press, New York.

Ogata, N., Alter, H. J., Miller, R. H., and Purcell, R. H. (1991). Nucleotide sequence and mutation rate of the H strain of hepatitis C virus. *Proc. Natl. Acad. Sci. USA* 88, 3392–3396.

Okamoto, H., Kojima, M., Okada, S., Yoshizawa, H., Iizuka, H., Tanaka, T., Muchmore, E. E., Peterson, D. A., Ito, Y., and Mishiro, S. (1992a). Genetic drift of hepatitis C virus during an 8.2-year infection in a chimpanzee: Variability and stability. *Virology* 190, 894–899.

Okamoto, H., Kurai, K., Okada, S., Yamamoto, K., Iizuka, H., Tanaka, T., Fukuda, S., Tsuda, F., and Mishiro, S. (1992b). Full-length sequence of a hepatitis C virus genome having poor homology to reported isolates: Comparative study of four distinct genotypes. *Virology* 188, 331–341.

Okamoto, H., Nakao, H., Inoue, T., Fukuda, M., Kishimoto, J., Iizuka, H., Tsuda, F., Miyakawa, Y., and Mayumi, M. (1997). The entire nucleotide sequences of two GB virus C/hepatitis G virus isolates of distinct genotypes from Japan. *J. Gen. Virol.* 78, 737–745.

Prince, A. M., Huima-Byron, T., Parker, T. S., and Levine, D. M. (1996). Visualization of hepatitis C virions and putative defective interfering particles isolated from low-density lipoproteins. *J. Viral Hepatitis* 3, 11–17.

Saitou, N., and Nei, M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sato, K., Okamoto, H., Aihara, S., Hoshi, Y., Tanaka, T., and Mishiro, S. (1993). Demonstration of sugar moiety on the surface of hepatitis C virions recovered from the circulation of infected humans. *Virology* 196, 354–357.

Sato, K., Tanaka, T., Okamoto, H., Miyakawa, Y., and Mayumi, M. (1996). Association of circulating hepatitis G virus with lipoproteins for a lack of binding with antibodies. *Biochem. Biophys. Res. Commun.* 229, 719–725.

Schmidt, B., Korn, K., and Fleckenstein, B. (1996). Molecular evidence for transmission of hepatitis G virus by blood transfusion. *Lancet* 347, 909. [Letter]

Schreiber, G. B., Busch, M. P., Kleinman, S. H., Korelitz, J. J., and the Retrovirus Epidemiology Donor Study. (1996). The risk of transfusion-transmitted viral infections. *N. Engl. J. Med.* 334, 1685–1690.

Shimizu, M., Osada, K., and Okamoto, H. (1996). Transfusion-transmitted hepatitis G virus following open heart surgery. *Transfusion* 36, 937. [Letter]

Simmonds, P. (1995). Variability of hepatitis C virus. *Hepatology* 21, 570–583.

Simons, J. N., Leary, T. P., Dawson, G. J., Pilot-Matias, T. J., Muerhoff, A. S., Schlauder, G. G., Desai, S. M., and Mushahwar, I. K. (1995). Isolation of novel virus-like sequences associated with human hepatitis. *Nature Med.* 1, 564–569.

Taniguchi, S., Okamoto, H., Sakamoto, M., Kojima, M., Tsuda, F., Tanaka, T., Munekata, E., Muchmore, E. E., Peterson, D. A., and Mishiro, S. (1993). A structurally flexible and antigenically variable N-terminal domain of the hepatitis C virus E2/NS1 protein: Implication for an escape from antibody. *Virology* 195, 297–301.

Thomssen, R., Bonk, S., Propfe, C., Heermann, K. H., Kochel, H. G., and Uy, A. (1992). Association of hepatitis C virus in human sera with beta-lipoprotein. *Med. Microbiol. Immunol. (Berlin)* 181, 293–300.

Tsuda, F., Hadiwandowo, S., Sawada, N., Fukuda, M., Tanaka, T., Okamoto, H., Miyakawa, Y., and Mayumi, M. (1996). Infection with GB virus C (GBV-C) in patients with chronic liver disease or on maintenance hemodialysis in Indonesia. *J. Med. Virol.* 49, 248–252.

Wang, Y., Chen, H. S., Fan, M. H., Liu, H. L., An, P., Sawada, N., Tanaka, T., Tsuda, F., and Okamoto, H. (1997). Infection with GB virus C and hepatitis C virus in hemodialysis patients and blood donors in Beijing. *J. Med. Virol.* 52, 26–30.

Weiner, A. J., Brauer, M. J., Rosenblatt, J., Richman, K. H., Tung, J., Crawford, K., Bonino, F., Saracco, G., Choo, Q. L., Houghton, M., and Han, J. H. (1991). Variable and hypervariable domains are found in the regions of HCV corresponding to the flavivirus envelope and NS1 proteins and the pestivirus envelope glycoproteins. *Virology* 180, 842–848.

Yoshiba, M., Okamoto, H., and Mishiro, S. (1995). Detection of the GBV-C hepatitis virus genome in serum from patients with fulminant hepatitis of unknown aetiology. *Lancet* 346, 1131–1132.